Supplementary Information

Supplementary Methods:

Protein production and purification: SasG $G5^2$ and E-$G5^2$ (WT and mutants) expression and purification procedures were as previously described (1, 2).

FRET labels: Tryptophan (E500W and I555W) and cysteine (E532C and E613C) residues were introduced into SasG $G5^2$ and E-$G5^2$ constructs by site-directed mutagenesis. Both, E-$G5^2$-E500W-E532C and $G5^2$-I555W-E613C were labelled with 5-((((2-iodoacetyl)amino)ethyl)amino)naphthalene-1-sulfonic acid (1,5-IAEDANS; Life Technologies) following the manufacturer's instructions as described previously (1).

Equilibrium studies Equilibrium unfolding of the proteins was studied by urea denaturation under standard conditions (phosphate-buffered saline, 25°C). Folding was followed by intrinsic tyrosine (WT, proline- and glycine-to-alanine mutants, excitation wavelength 276 nm; emission 305 nm) and tryptophan (Y265W; excitation 280 nm; emission 350 nm) fluorescence and FRET measurements (excitation 280 nm: emission 490 nm) on a fluorescence spectrometer (Perkin Elmer LS55). The data were analyzed as previously described (1).

Kinetic studies: Kinetic experiments following the change in the fluorescence signal at different urea concentrations were carried out using a stopped-flow fluorimeter (Applied Photophysics SX.20) at 25°C constant temperature, as described previously (1). The data were fitted to equations describing single- or double-exponential phases (see text). To account for non-linearity in the observed unfolding rate constant, the chevron plot data were fitted to a sequential transition states model as described previously (3), in which denaturant induces a switch between two barriers separated by a high-energy intermediate.

Φ-values were determined using the following equation:

$$\Phi = \frac{\Delta\Delta G_{\text{D-}\ddagger}}{\Delta\Delta G_{\text{D-N}}}$$

Where $\Delta\Delta G_{\text{D-N}}$ was determined using equilibrium experiments, and $\Delta\Delta G_{\text{D-}\ddagger} = RT \ln\left(\frac{k_{\text{wt}}^{\text{H2O}}}{k_{\text{mut}}^{\text{H2O}}}\right)$.

Native reference structures were the crystal structure of *S. aureus* SasG E-$G5^2$ (PDB accession: 3TIP) for both E-$G5^2$ and the $G5^2$ domain alone.

Simulations: Simulations were performed using a coarse-grained model where only C_α atoms are represented and interactions depend on the native reference structure and on the residue type (4).

Equilibrium simulations were performed at a broad range of temperatures between 270 and 330 K lasting at least 30 μs. Temperature was controlled using Langevin dynamics, and the timestep for integration of the equations of motion was 15 fs. For both systems the mid-point temperature was approximately 320K. At this temperature E-G5^2 completely unfolds only a few times, hence we performed 62 simulations starting from random conformations sampled at 350K and setting the thermostat to temperatures between 270 and 315K at which E-G5^2 is expected to be folded. For all simulations in which full folding occurs, the pathway is identical to those observed during the equilibrium simulation reported in Fig. 3.

Crystallisation of E-G5^2-Y625W: E-G5^2 Y625W was purified as described previously (5) and concentrated to 47.6 mg.ml^{-1} in 20 mM Tris, 150 mM NaCl, pH 8. Crystallisation screening with JCSG+ (Molecular Dimensions; (6)) resulted in growth of large single crystals in conditions comprising 100 mM citrate pH 5 and 20% PEG 6000. Crystals were flash cooled in liquid N_2 prior to data collection on Diamond beamline I02. Data were indexed, integrated and scaled using XDS (7) and merged using Aimless (8). Phases were determined by molecular replacement with PhaserMR (9) using WT E-G5^2 (PDB accession: 3TIP (5)). E-G5^2-Y625W crystallised in spacegroup *C*2 with one molecule in the asymmetric unit. The model was improved using Coot (10) and refined to 1.6 Å (Supplementary Table 7) with nine translation/libration/screw (TLS) groups by Phenix (11). The coordinates and structure factors have been deposited in the protein data bank with accession code 5DBL. The structure was aligned by secondary structure matching with WT E-G5^2 using Superpose (12) and cartoons were rendered with CCP4mg (13).

Supplementary Figures:

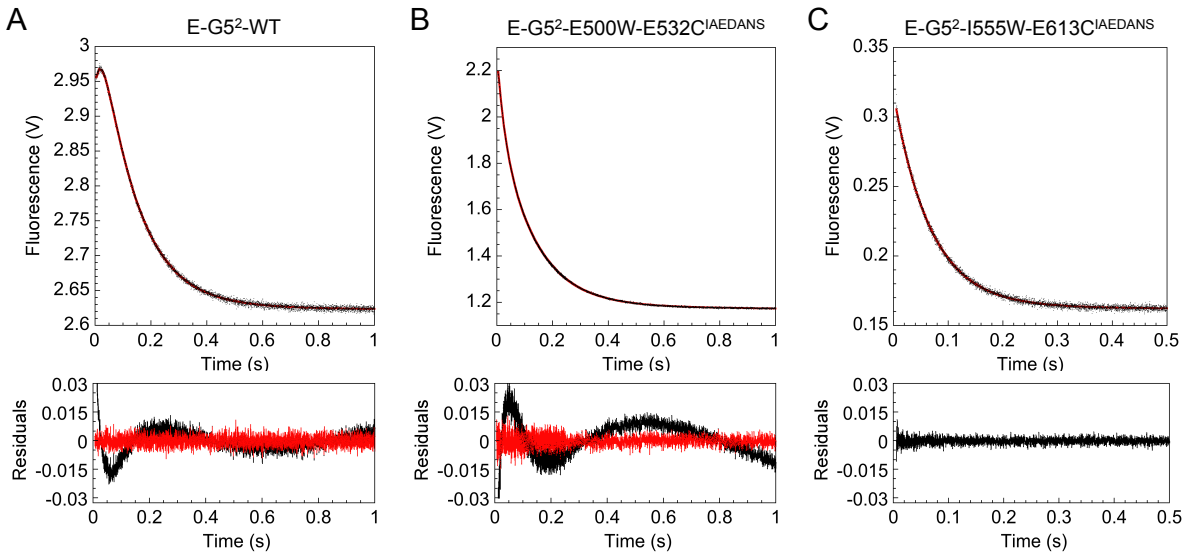

Fig. S1. At high denaturant concentrations two unfolding phases are observed in E-G5^2, as unfolding of the domains becomes uncoupled. Kinetics of unfolding into 9.5 M urea for E-G5^2-WT (*A*), E-G5^2-E500W-E532C^{IAEDANS} (*B*) and E-G5^2-I555W-E613C^{IAEDANS} (*C*). Traces were collected by monitoring the change in intrinsic tyrosine or 1,5-IAEDANS fluorescence. Unfolding traces of E-G5^2-WT (*A*) and E-G5^2-E500W-E532C^{IAEDANS} (*B*) were fitted to the sum of two exponentials, which describes the data better than the single exponential. Residuals for the fit to the single exponential and the sum of two exponentials are shown below the data in black and red, respectively. Unfolding traces of E-G5^2-I555W-E613C^{IAEDANS} (*C*) (that monitors the unfolding of the G5 domain only) were fitted to a single exponential, which describes the data well. The residuals are shown below the data.

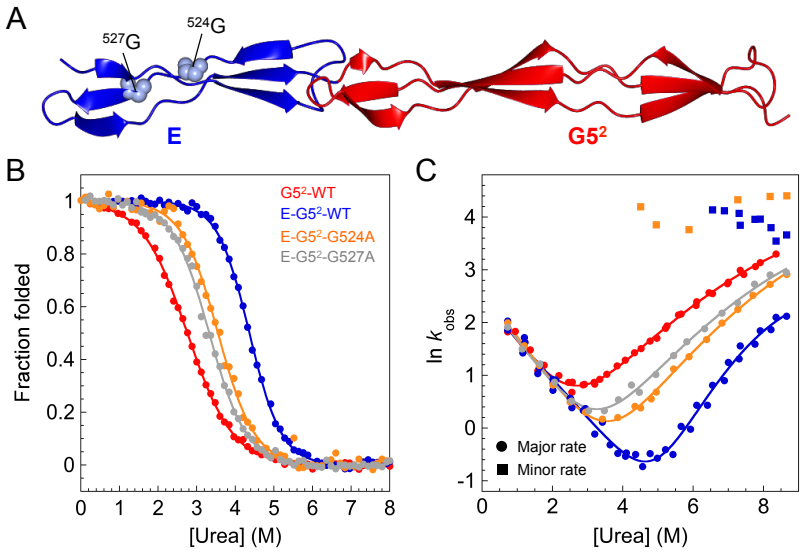

Fig. S2. Highly destabilizing mutations in the E-domain break the cooperative unfolding of E-G5^2. (*A*) Structure of E-G5^2 showing the location of mutated residues within the E domain (Gly524, Gly527 light blue spheres) (*B*) Equilibrium denaturation curves and (*C*) urea dependence of the natural logarithm of the observed rate constants for wild type and mutant proteins. Circles and squares represent major and minor rate constants, respectively. Note that the unfolding *m*-value of these two mutants reverts to that of wild-type G5^2 showing that the E and G5^2 domains are now unfolding independently.

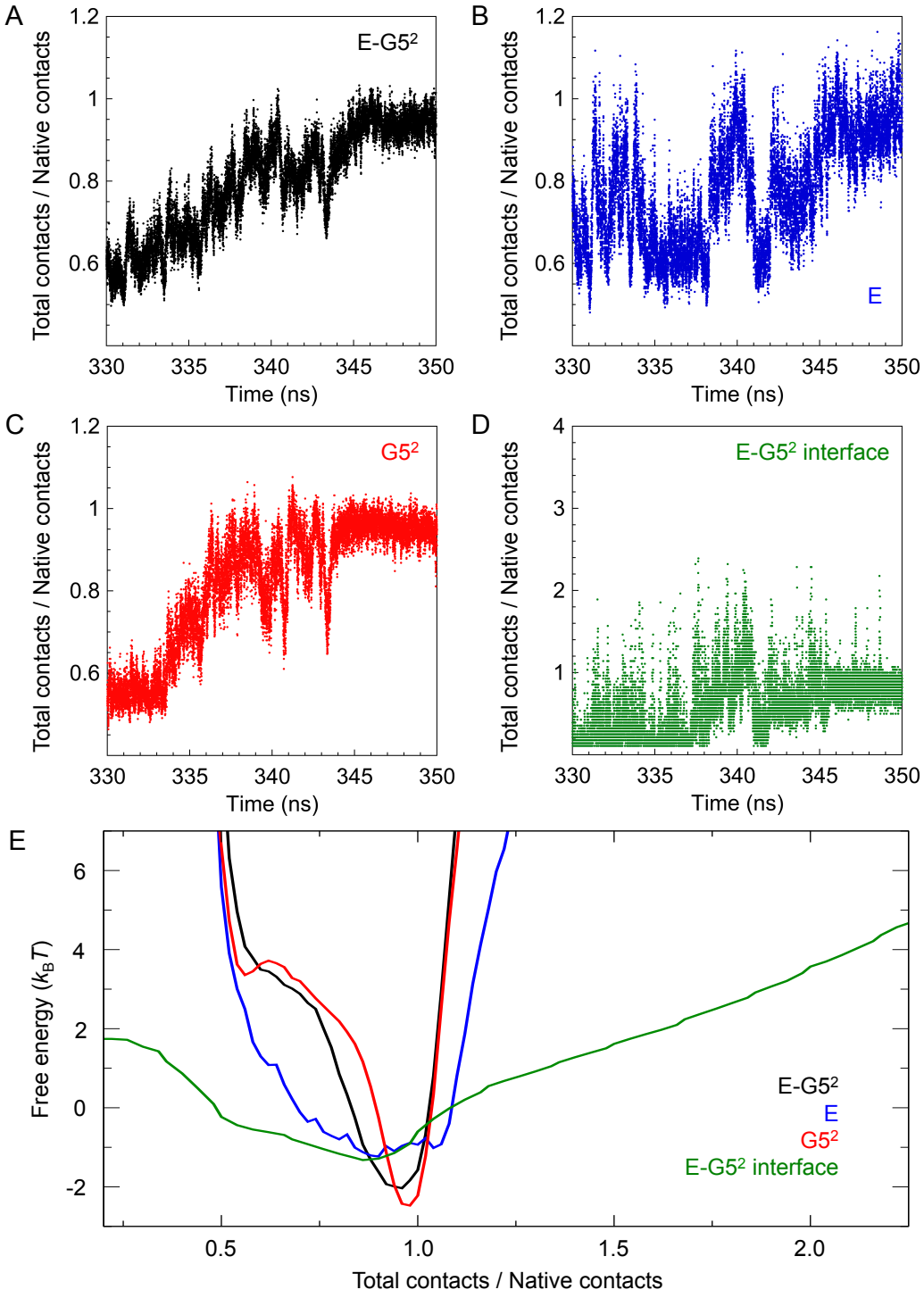

Fig. S3. Simulations of E-G5^2 at 320 K. Trajectories of the total contacts normalized by the number of native contacts for E-G5^2 (*A*), E (*B*), G5^2 (*C*) and the E-G5^2 interface (*D*), for the same folding event as presented in Fig. 3. Panel *E* shows the free energy change as a function of the ratio of total contacts to native contacts. Domain E is characterized by a broad basin that encompasses both folded and unfolded states whereas the G5^2 domain shows a barrier between the unfolded and folded state (at 320 K the native state is much more populated that the unfolded state).

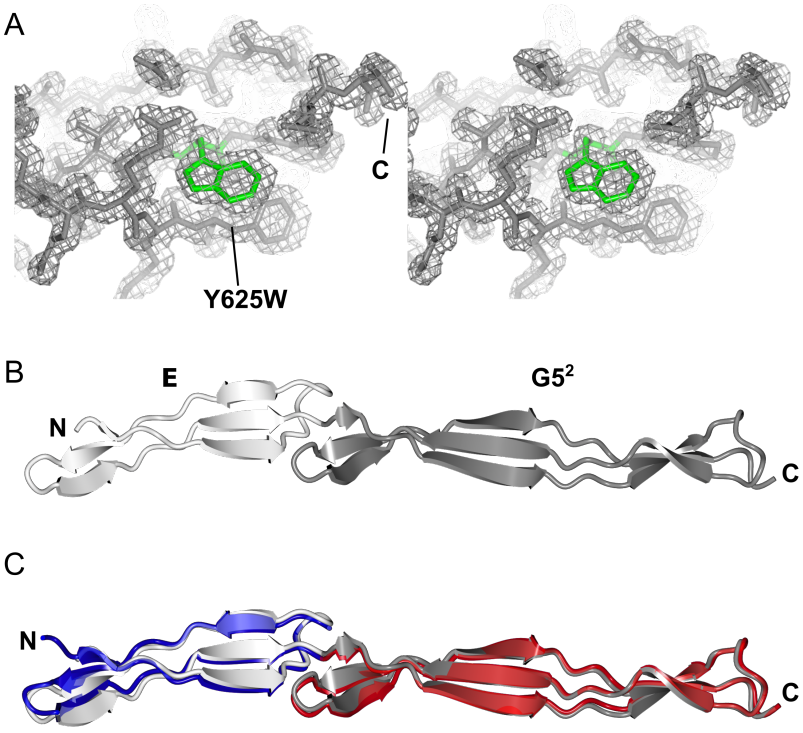

Fig. S4. The structure of E-G5^2-Y625W is highly similar to the wild type protein fold. An X-ray crystal structure of E-G5^2-Y625W was determined at 1.6 Å resolution (PDB: 5DBL). (*A*) Stereo image of the *2mFo-DFc* electron density map (grey) contoured at 1 electron/Å^3 at the C-terminus of G5^2; the Y625W side-chain is shown in green. (*B*) The X-ray crystal structure of E-G5^2 Y625W (E, white and G5^2, grey) is highly similar to the wild type (PDB accession: 3TIP E, blue and G5^2, red). Alignment by secondary structure matching revealed a Cα root mean square deviation of 1 Å (*C*), confirming the Y625W mutation does not affect the overall structure of E-G5^2.

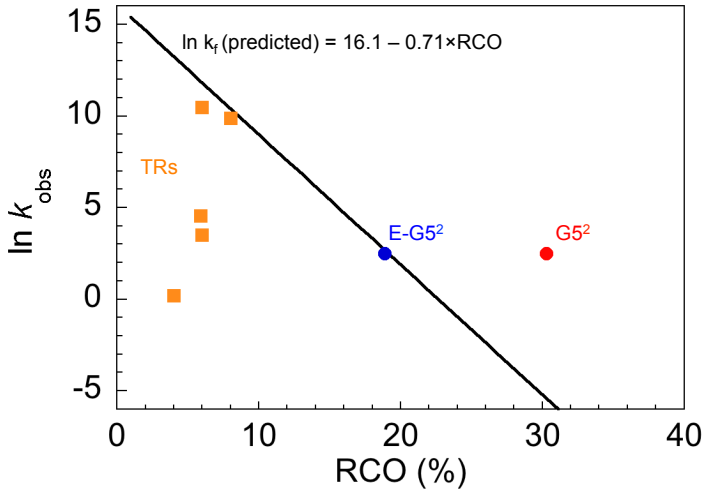

Fig. S5. Plot of rate of folding vs contact order. $G5^2$ folds significantly more rapidly than would be predicted from its relative contact order. (Data from Plaxco et al (14) shown by the straight line). Although it has some properties of a repeat protein SasG clearly lacks the short-range interactions that characterize all true tandem repeat proteins (TRs; examples include leucine-rich repeats, ankyrin repeats, and tetratricopeptide repeats, orange).

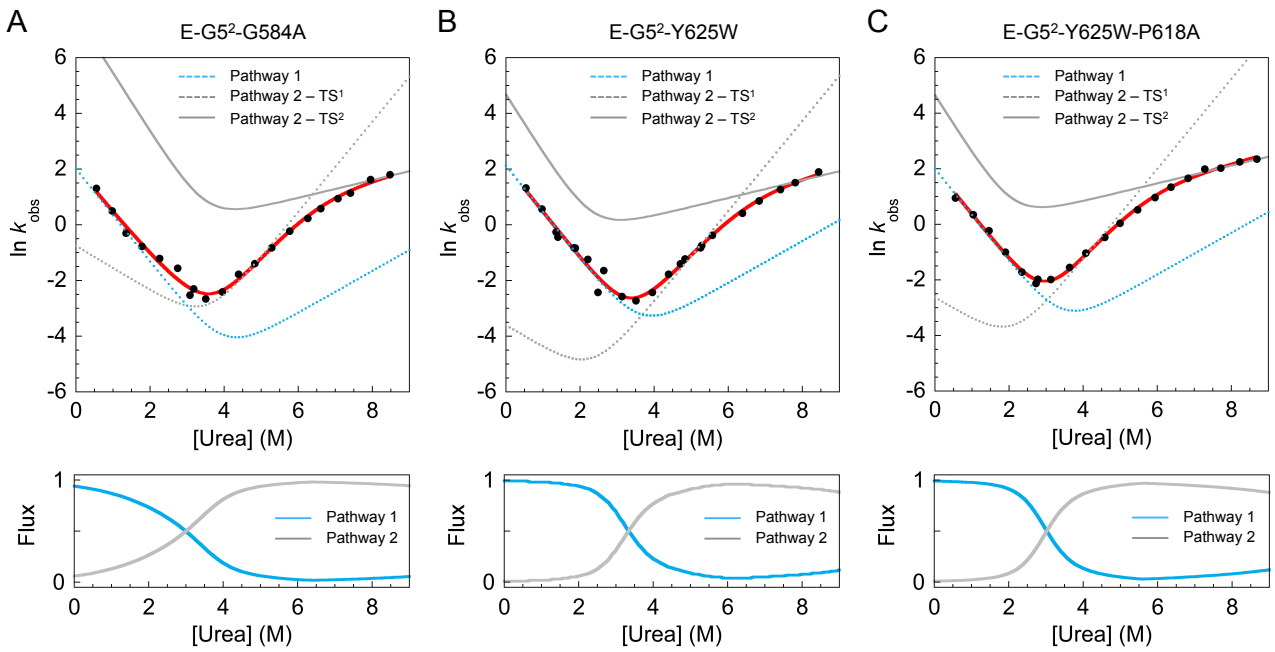

Fig. S6. Fitting the data to parallel pathways models. Chevron and flux plots for representative E-G5^2 mutants that fold via an alternative pathway: E-G5^2-G584A (*A*), E-G5^2-Y625W (*B*) and E-G5^2-Y625W-P618A (*C*). The chevron plots were fitted globally to a model assuming two parallel pathways, shown in red, in which the observed rate constant is equal to the sum of the rate constants for each pathway (for details see Table S6). The hypothetical chevron corresponding to the alternative pathway (pathway 1) is shown as a dashed blue line. The other (wild-type) pathway (pathway 2) was assumed to follow the sequential transition states model (3) as the wild-type pathway. The hypothetical chevrons corresponding to the wild-type pathway transition state 1 (TS^1) and transition state 2 (TS^2) are shown as dashed orange and purple lines, respectively. The bottom plots illustrate the fractional fluxes through the alternative pathway (pathway 1, blue) and wild-type pathway (pathway 2, grey).

Supplementary Tables:

Table S1. Thermodynamic and kinetic parameters for $G5^2$ and E-$G5^2$.

Protein	Equilibrium*		Kinetic†			
	m_{D-N}* (kcal·mol^{-1}·M^{-1})	$\Delta G_{D-N}^{H_2O}$* (kcal·mol^{-1})	m_{D-N}† (kcal·mol^{-1}·M^{-1})	$\Delta G_{D-N}^{H_2O}$† (kcal·mol^{-1})	$k_f^{H_2O}$† (s^{-1})	$k_u^{H_2O}$† (s^{-1})
$G5^2$-WT	1.0 ± 0.1	2.8 ± 0.2	1.1 ± 0.1	2.8 ± 0.2	12.2 ± 0.3	0.112 ± 0.025
E-$G5^2$-WT	1.4 ± 0.1	6.3 ± 0.2	1.4 ± 0.1	7.0 ± 0.3	13.0 ± 0.3	(9.4 ± 3.2)×10^{-5}
E-$G5^2$-Y547	1.4 ± 0.1	5.5 ± 0.1	-	-	-	-
E-$G5^2$-Y625	1.4 ± 0.1	5.6 ± 0.1	1.4 ± 0.1	6.5 ± 0.4	10.6 ± 0.5	(19.5 ± 6.9)×10^{-5}
$EG5^2$-T501C^{A488} -E613C^{A594}	1.4 ± 0.1	6.1 ± 0.3	1.4 ± 0.1	6.7 ± 0.4	10.7 ± 0.5	(14.0 ± 5.0)×10^{-5}
$EG5^2$-E500W-E532C^{IAEDANS}	1.4 ± 0.1	6.1 ± 0.4	1.4 ± 0.1	6.7 ± 0.4	10.3 ± 0.5	(11.8 ± 4.2)×10^{-5}
$EG5^2$-E555I-E613C^{IAEDANS}	1.4 ± 0.1	4.5 ± 0.3	1.4 ± 0.1	5.4 ± 0.3	14.0 ± 0.7	(1.5 ± 0.5)×10^{-3}

* Equilibrium parameters were obtained by fitting the data to a two-state equation.

† Kinetic parameters were calculated from fitting the data globally to a sequential transition states model.

Table S2. Apparent equilibrium parameters obtained for wild-type G5[2] and its mutants at 25°C.

Protein	$m_{\mathrm{D-N}}$ (kcal·mol^{-1}·M^{-1})	$[D]_{50\%}$ (M^{-1})	$\Delta G_{\mathrm{D-N}}^{\mathrm{H_2O}}$ (kcal·mol^{-1})	$\Delta\Delta G_{\mathrm{D-N}}^{\mathrm{H_2O}}$ (kcal·mol^{-1})
G5[2]-WT	1.00 ± 0.05	2.80 ± 0.07	2.80 ± 0.16	-
G5[2]-P549A	0.91 ± 0.02	2.42 ± 0.03	2.20 ± 0.05	0.60 ± 0.16
G5[2]-P562A	1.02 ± 0.02	2.22 ± 0.01	2.26 ± 0.05	0.54 ± 0.17
G5[2]-P571A	0.98 ± 0.02	1.86 ± 0.05	1.83 ± 0.06	0.97 ± 0.17
G5[2]-P575A	1.03 ± 0.02	1.94 ± 0.01	1.99 ± 0.04	0.81 ± 0.16
G5[2]-P594A	0.99 ± 0.02	2.40 ± 0.01	2.37 ± 0.04	0.43 ± 0.16
G5[2]-P599A	0.99 ± 0.02	2.70 ± 0.01	2.68 ± 0.04	0.12 ± 0.16
G5[2]-P618A	1.03 ± 0.02	2.18 ± 0.01	2.24 ± 0.05	0.56 ± 0.16
G5[2]-P627A	0.96 ± 0.03	1.67 ± 0.02	1.62 ± 0.05	1.18 ± 0.16
G5[2]-G548A	0.92 ± 0.04	2.94 ± 0.06	2.71 ± 0.14	0.09 ± 0.21
G5[2]-G552A	1.05 ± 0.05	2.90 ± 0.05	3.03 ± 0.16	-0.23 ± 0.22
G5[2]-G576A	1.01 ± 0.04	0.85 ± 0.19	0.85 ± 0.19	1.95 ± 0.25
G5[2]-G584A	1.00	-1.60 ± 0.61	-1.6 ± 0.61	4.40 ± 0.63
G5[2]-G587A	-	-	-	-
G5[2]-G602A	1.06 ± 0.04	1.28 ± 0.03	1.35 ± 0.07	1.45 ± 0.17
G5[2]-G608A	1.02 ± 0.03	1.58 ± 0.02	1.61 ± 0.05	1.19 ± 0.17
G5[2]-G626A	0.98 ± 0.07	1.11 ± 0.25	1.08 ± 0.25	1.72 ± 0.30
G5[2]-Y625W	0.92 ± 0.02	1.53 ± 0.01	1.41 ± 0.03	1.39 ± 0.16

The parameters were calculated by fitting the equilibrium denaturation curves to a two-state model. The errors quoted for $m_{\mathrm{D-N}}$ and $[D]_{50\%}$ of G5[2]-WT represent the experimental errors (based on four independent experiments). The errors quoted for the G5[2] mutants are the errors of the fits of the data. In the case of G5[2]-G584A, the data were fit to a two-state equation with the $m_{\mathrm{D-N}}$ value fixed at 1 kcal·mol^{-1}·M^{-1}. G5[2]-G587A is inherently unstable in water.

Table S3. Apparent equilibrium parameters obtained for wild-type E-G5[2] and its mutants at 25°C.

Protein	$m_{\mathrm{D-N}}$ (kcal·mol^{-1}·M^{-1})	$[D]_{50\%}$ (M^{-1})	$\Delta G_{\mathrm{D-N}}^{\mathrm{H_2O}}$ (kcal·mol^{-1})	$\Delta\Delta G_{\mathrm{D-N}}^{\mathrm{H_2O}}$ (kcal·mol^{-1})
E-G5[2]-WT	1.42 ± 0.04	4.40 ± 0.02	6.27 ± 0.18	-
E-G5[2]-Y547	1.40 ± 0.02	3.92 ± 0.03	5.49 ± 0.09	0.78 ± 0.20
E-G5[2]-Y625	1.37 ± 0.03	4.07 ± 0.01	5.58 ± 0.10	0.69 ± 0.21
EG5[2]-T501C^{A488}-E613C^{A594}	1.40 ± 0.06	4.35 ± 0.02	6.09 ± 0.28	0.18 ± 0.33
EG5[2]-E500W-E532C^{IAEDANS}	1.39 ± 0.06	4.38 ± 0.02	6.10 ± 0.27	0.17 ± 0.33
EG5[2]-E555I-E613C^{IAEDANS}	1.44 ± 0.06	3.13 ± 0.02	4.50 ± 0.18	1.77 ± 0.26
E-G5[2]-P499A	1.38 ± 0.06	4.41 ± 0.02	6.10 ± 0.26	0.17 ± 0.31
E-G5[2]-P504A	1.34 ± 0.04	3.99 ± 0.01	5.36 ± 0.14	0.91 ± 0.23
E-G5[2]-P512A	1.28 ± 0.03	4.08 ± 0.01	5.22 ± 0.12	1.05 ± 0.22
E-G5[2]-P515A	1.40 ± 0.04	4.26 ± 0.01	5.97 ± 0.16	0.29 ± 0.24
E-G5[2]-P523A	1.32 ± 0.03	4.10 ± 0.01	5.40 ± 0.14	0.86 ± 0.23
E-G5[2]-P526A	1.30 ± 0.04	4.13 ± 0.02	5.37 ± 0.18	0.90 ± 0.26
E-G5[2]-P531A	1.35 ± 0.04	4.16 ± 0.02	5.63 ± 0.17	0.64 ± 0.25
E-G5[2]-P539A	1.37 ± 0.06	4.38 ± 0.03	6.02 ± 0.28	0.25 ± 0.33
E-G5[2]-P540A	1.40 ± 0.10	4.55 ± 0.04	6.37 ± 0.45	-0.10 ± 0.48
E-G5[2]-P549A	1.32 ± 0.04	3.68 ± 0.02	4.85 ± 0.15	1.42 ± 0.24
E-G5[2]-P562A	1.44 ± 0.04	3.90 ± 0.01	5.61 ± 0.14	0.65 ± 0.23
E-G5[2]-P571A	1.33 ± 0.04	3.92 ± 0.02	5.19 ± 0.15	1.07 ± 0.24
E-G5[2]-P575A	1.41 ± 0.04	3.99 ± 0.02	5.63 ± 0.18	0.63 ± 0.26
E-G5[2]-P594A	1.28 ± 0.04	4.19 ± 0.02	5.39 ± 0.17	0.88 ± 0.25
E-G5[2]-P599A	1.17 ± 0.03	2.94 ± 0.02	3.44 ± 0.10	2.83 ± 0.21
E-G5[2]-P599A-E500W-E532C^{IAEDANS}	0.93 ± 0.02	2.03 ± 0.01	1.88 ± 0.04	4.39 ± 0.19
E-G5[2]-P618A	1.32 ± 0.06	3.84 ± 0.02	5.07 ± 0.21	1.20 ± 0.28
E-G5[2]-P627A	1.28 ± 0.04	3.77 ± 0.02	4.83 ± 0.14	1.44 ± 0.23
E-G5[2]-G505A	1.34 ± 0.07	4.07 ± 0.03	5.45 ± 0.30	0.82 ± 0.36
E-G5[2]-G517A	1.35 ± 0.05	3.38 ± 0.02	4.58 ± 0.19	1.69 ± 0.26
E-G5[2]-G524A	1.23 ± 0.06	3.58 ± 0.03	4.40 ± 0.22	1.87 ± 0.28
E-G5[2]-G527A	1.23 ± 0.04	3.37 ± 0.02	4.16 ± 0.15	2.10 ± 0.24

E-G5^2-G534A	1.23 ± 0.03	3.34 ± 0.01	4.11 ± 0.09	2.16 ± 0.21
E-G5^2-G548A	1.20 ± 0.06	3.30 ± 0.03	3.97 ± 0.19	2.30 ± 0.26
E-G5^2-G552A	1.30 ± 0.04	3.66 ± 0.02	4.75 ± 0.15	1.52 ± 0.24
E-G5^2-G576A	1.48 ± 0.05	3.18 ± 0.02	4.72 ± 0.16	1.55 ± 0.24
E-G5^2-G584A	1.45 ± 0.05	3.20 ± 0.02	4.62 ± 0.15	1.64 ± 0.23
E-G5^2-G587A	1.51 ± 0.05	1.25 ± 0.02	1.88 ± 0.07	4.39 ± 0.19
E-G5^2-G602A	1.41 ± 0.03	3.04 ± 0.01	4.29 ± 0.09	1.97 ± 0.21
E-G5^2-G608A	1.41 ± 0.04	3.47 ± 0.02	4.90 ± 0.14	1.37 ± 0.23
E-G5^2-G626A	1.40 ± 0.05	3.01 ± 0.02	4.23 ± 0.14	2.04 ± 0.23
E-G5^2-Y625W	1.53 ± 0.02	3.37 ± 0.01	5.14 ± 0.08	1.13 ± 0.20
E-G5^2-Y625W-P512A	1.48 ± 0.03	3.16 ± 0.01	4.67 ± 0.11	1.60 ± 0.21
E-G5^2-Y625W-P531A	1.51 ± 0.02	3.41 ± 0.01	5.13 ± 0.05	1.14 ± 0.19
E-G5^2-Y625W-P540A	1.63 ± 0.04	3.54 ± 0.01	5.78 ± 0.14	0.49 ± 0.23
E-G5^2-Y625W-P571A	1.51 ± 0.04	3.10 ± 0.02	4.67 ± 0.14	1.60 ± 0.23
E-G5^2-Y625W-P599A	1.42 ± 0.02	1.96 ± 0.01	2.79 ± 0.04	3.48 ± 0.19
E-G5^2-Y625W-P618A	1.57 ± 0.02	3.03 ± 0.01	4.75 ± 0.07	1.52 ± 0.19
E-G5^2-G584A-E500W-E532C^{IAEDANS}	1.58 ± 0.09	2.01 ± 0.04	3.17 ± 0.18	3.10 ± 0.26

The parameters were calculated by fitting the equilibrium denaturation curves to a two-state model. The errors quoted for the G5^2 mutants are the errors of the fits of the data.

Table S4. Kinetic parameters obtained for the $G5^2$ pathway based on the single mutants at 25°C.

Protein	Equilibrium $\Delta\Delta G_{D-N}^{H_2O}$ (kcal·mol^{-1})	Kinetic $\Delta\Delta G_{D-N}^{H_2O}$ (kcal·mol^{-1})	$k_f^{H_2O}$ (s^{-1})	$k_u^{H_2O}$ (s^{-1})	Φ
$G5^2$-WT	-	-	12.2 ± 0.3	0.11 ± 0.02	-
$G5^2$-P549A	0.6 ± 0.2	-0.2 ± 0.3	13.9 ± 0.5	0.09 ± 0.02	-
$G5^2$-P562A	0.5 ± 0.2	0.6 ± 0.3	13.0 ± 0.5	0.30 ± 0.07	-0.07
$G5^2$-P571A	1.0 ± 0.2	0.6 ± 0.3	5.1 ± 0.2	0.13 ± 0.03	0.91
$G5^2$-P575A	0.8 ± 0.2	0.6 ± 0.3	5.0 ± 0.2	0.13 ± 0.03	0.90
$G5^2$-P594A	0.4 ± 0.2	0.2 ± 0.3	13.9 ± 0.5	0.18 ± 0.04	-
$G5^2$-P599A	0.1 ± 0.2	0.0 ± 0.3	12.4 ± 0.5	0.11 ± 0.03	-
$G5^2$-P618A	0.6 ± 0.2	0.7 ± 0.3	13.5 ± 0.6	0.42 ± 0.09	-0.11
$G5^2$-P627A	1.2 ± 0.2	0.9 ± 0.3	2.5 ± 0.1	0.11 ± 0.02	1.02
$G5^2$-G548A	0.1 ± 0.2	-0.4 ± 0.3	13.8 ± 0.5	0.07 ± 0.02	-
$G5^2$-G552A	-0.2 ± 0.2	-0.3 ± 0.3	14.1 ± 0.5	0.08 ± 0.02	-
$G5^2$-G576A	2.0 ± 0.2	1.3 ± 0.3	0.9 ± 0.1	0.07 ± 0.02	0.80
$G5^2$-G602A	1.5 ± 0.2	1.6 ± 0.3	8.5 ± 0.4	1.16 ± 0.30	0.05
$G5^2$-G608A	1.2 ± 0.2	1.1 ± 0.3	13.3 ± 0.6	0.82 ± 0.18	0.01
$G5^2$-G626A	1.7 ± 0.3	2.0 ± 0.3	0.6 ± 0.1	0.15 ± 0.03	1.06
$G5^2$-Y625W	1.4 ± 0.2	1.5 ± 0.3	1.3 ± 0.1	0.14 ± 0.03	0.95

The chevron plots were fitted globally to the sequential transition states model, with the values of k_{I^*-D} and m_{I^*-D} fixed at 1×10^4 s^{-1} and 0 M^{-1}, respectively, and the values of m_{D-I^*}, m_{I^*-N} and m_{N-I^*} shared between the data sets (0.88 ± 0.02 M^{-1}, 0.64 ± 0.02 M^{-1} and 0.29 ± 0.02 M^{-1}, respectively; kinetic m_{D-N} was 1.07 ± 0.03 kcal·mol^{-1}·M^{-1}). All other microscopic rate constants were allowed to vary freely. Φ values were calculated using equilibrium rather than kinetic $\Delta\Delta G_{D-N}^{H_2O}$, due to lower associated errors. The rate constants and Φ values presented in the table are for TS[1]. The errors on the Φ values are 5-10%.

Table S5. Kinetic parameters obtained for the main (wild-type) E-G5[2] pathway at 25°C.

Protein	Equilibrium $\Delta\Delta G_{\mathrm{D-N}}^{\mathrm{H_2O}}$ (kcal·mol^{-1})	Kinetic $\Delta\Delta G_{\mathrm{D-N}}^{\mathrm{H_2O}}$ (kcal·mol^{-1})	$k_{\mathrm{f}}^{\mathrm{H_2O}}$ (s^{-1})	$k_{\mathrm{u}}^{\mathrm{H_2O}}$ (s^{-1})	Φ
E-G5[2]-WT	-	-	13.0 ± 0.3	(9.4 ± 3.2)×10^{-5}	-
E-G5[2]-Y625	0.7 ± 0.2	0.6 ± 0.5	10.6 ± 0.5	(2.0 ± 0.7)×10^{-4}	-
EG5[2]-T501C^{A488}-E613C^{A594}	0.2 ± 0.3	0.4 ± 0.5	10.7 ± 0.5	(1.4 ± 0.5)×10^{-4}	-
EG5[2]-E500W-E532C^{IAEDANS}	0.2 ± 0.3	0.3 ± 0.5	10.3 ± 0.5	(1.2 ± 0.4)×10^{-4}	-
EG5[2]-E555I-E613C^{IAEDANS}	1.8 ± 0.3	1.6 ± 0.5	14.0 ± 0.7	(1.5 ± 0.5)×10^{-3}	-
E-G5[2]-P499A	0.2 ± 0.3	0.3 ± 0.5	11.1 ± 0.5	(1.3 ± 0.5)×10^{-4}	-
E-G5[2]-P504A	0.9 ± 0.2	1.1 ± 0.5	10.1 ± 0.5	(4.3 ± 1.5)×10^{-4}	0.17
E-G5[2]-P512A	1.0 ± 0.2	0.7 ± 0.5	12.6 ± 0.5	(2.8 ± 1.0)×10^{-4}	0.02
E-G5[2]-P515A	0.3 ± 0.2	0.4 ± 0.5	11.6 ± 0.5	(1.7 ± 0.6)×10^{-4}	-
E-G5[2]-P523A	0.9 ± 0.2	0.8 ± 0.5	11.2 ± 0.5	(2.9 ± 1.0)×10^{-4}	0.10
E-G5[2]-P526A	0.9 ± 0.3	0.5 ± 0.5	11.1 ± 0.5	(1.8 ± 0.6)×10^{-4}	0.10
E-G5[2]-P531A	0.6 ± 0.3	0.5 ± 0.5	11.9 ± 0.5	(1.9 ± 0.7)×10^{-4}	-
E-G5[2]-P539A	0.3 ± 0.3	0.0 ± 0.5	12.2 ± 0.5	(0.9 ± 0.4)×10^{-4}	-
E-G5[2]-P540A	-0.1 ± 0.5	0.0 ± 0.5	11.5 ± 0.5	(0.9 ± 0.3)×10^{-4}	-
E-G5[2]-P549A	1.4 ± 0.2	1.1 ± 0.5	12.1 ± 0.6	(5.1 ± 1.8)×10^{-4}	0.03
E-G5[2]-P562A	0.7 ± 0.2	0.5 ± 0.5	12.8 ± 0.6	(5.1 ± 1.8)×10^{-4}	0.02
E-G5[2]-P571A	1.1 ± 0.2	0.4 ± 0.5	5.8 ± 0.3	(9.0 ± 3.2)×10^{-5}	1.03
E-G5[2]-P575A	0.6 ± 0.3	0.6 ± 0.5	5.1 ± 0.2	(9.7 ± 3.4)×10^{-5}	0.97
E-G5[2]-P594A	0.9 ± 0.3	0.5 ± 0.5	13.3 ± 0.6	(2.1 ± 0.7)×10^{-4}	-0.01
E-G5[2]-P618A	1.2 ± 0.3	0.5 ± 0.5	13.0 ± 0.6	(2.2 ± 0.8)×10^{-4}	0.00
E-G5[2]-P627A	1.4 ± 0.2	0.8 ± 0.5	3.3 ± 0.1	(9.8 ± 3.3)×10^{-5}	0.98
E-G5[2]-G505A	0.8 ± 0.4	0.6 ± 0.5	13.1 ± 0.6	(2.4 ± 0.8)×10^{-4}	0.00
E-G5[2]-G534A	2.2 ± 0.2	0.5 ± 0.5	13.0 ± 0.6	(2.4 ± 0.8)×10^{-4}	0.00
E-G5[2]-G602A	2.0 ± 0.2	1.9 ± 0.5	9.9 ± 0.5	(1.7 ± 0.6)×10^{-3}	0.08
E-G5[2]-G608A	1.4 ± 0.2	1.2 ± 0.5	12.1 ± 0.6	(6.9 ± 2.4)×10^{-4}	0.03
E-G5[2]-Y625W-P599A	3.5 ± 0.2	3.9 ± 0.5	1.2 ± 0.1	(6.6 ± 1.9)×10^{-3}	-

The chevron plots were fitted globally to the sequential transition states model, with the values of $k_{\mathrm{I^*-D}}$ and $m_{\mathrm{I^*-D}}$ fixed at 1×10^4 s^{-1} and 0 M^{-1}, respectively, and the values of $m_{\mathrm{D-I^*}}$, $m_{\mathrm{I^*-N}}$ and $m_{\mathrm{N-I^*}}$ shared between the data sets (0.80 ± 0.01 M^{-1}, 1.30 ± 0.04 M^{-1} and 0.32 ± 0.03 M^{-1}, respectively; kinetic $m_{\mathrm{D-N}}$ was 1.43 ± 0.05 kcal·mol^{-1}·M^{-1}). All other microscopic rate constants were allowed to vary freely. Φ values were calculated using equilibrium rather than kinetic $\Delta\Delta G_{\mathrm{D-N}}^{\mathrm{H_2O}}$, due to lower associated errors. The rate constants and Φ values presented in the table are for TS[1]. The errors on the Φ values are 5-10%. Due to little confidence in the Φ values calculated for TS[2] (errors of 5-40%, owing to large errors in the rate constants associated with TS[2]), they are not listed.

Table S6. Kinetic parameters for the alternative folding pathway of E-G5[2] at 25°C obtained from the parallel pathways model fitting.

Protein	Pathway 1		Pathway 2			
			TS[1]		TS[2]	
	$k_f^{H_2O}$ (s^{-1})	$k_u^{H_2O}$ (s^{-1})	$k_f^{H_2O}$ (s^{-1})	$k_u^{H_2O}$ (s^{-1})	$k_f^{H_2O}$ (s^{-1})	$k_u^{H_2O}$ (s^{-1})
E-G5[2]-G576A	8.4 ± 1.3	(1.4 ± 1.7)×10^{-3}	0.03 ± 0.27	(9.8 ± 3.4)×10^{-5}	(0.1 ± 1.0)×10^{3}	0.38 ± 0.07
E-G5[2]-G584A	7.5 ± 1.2	(0.5 ± 1.4)×10^{-3}	0.47 ± 0.25	(9.6 ± 2.8)×10^{-5}	(1.9 ± 1.1)×10^{3}	0.39 ± 0.05
E-G5[2]-Y625W	7.9 ± 0.6	(1.4 ± 0.4)×10^{-3}	0.066 ± 0.088	(1.9 ± 0.2)×10^{-4}	(1.8 ± 2.5)×10^{2}	0.54 ± 0.03
E-G5[2]-Y625W P512A	6.0 ± 0.7	(0.8 ± 2.3)×10^{-3}	0.14 ± 0.20	(3.5 ± 0.7)×10^{-4}	(2.1 ± 2.9)×10^{2}	0.51 ± 0.07
E-G5[2]-Y625W P531A	7.5 ± 0.6	(6.7 ± 8.7)×10^{-4}	0.87 ± 0.16	(1.2 ± 0.2)×10^{-4}	(2.8 ± 0.6)×10^{3}	0.39 ± 0.03
E-G5[2]-Y625W P540A	15.3 ± 2.1	(6.7 ± 8.7)×10^{-4}	0.27 ± 0.36	(1.1 ± 0.3)×10^{-4}	(0.1 ± 1.3)×10^{3}	0.38 ± 0.06
E-G5[2]-Y625W -P618A	7.4 ± 0.7	(1.8 ± 1.8)×10^{-3}	0.07 ± 0.17	(4.4 ± 0.6)×10^{-4}	(1.0 ± 2.4)×10^{2}	0.64 ± 0.06

The chevron plots were fitted globally to a model assuming two parallel pathways, in which the observed rate constant is equal to the sum of the rate constants for each pathway (see Fig. S5 for representative examples). For the alternative folding pathway (pathway 1) we assumed the simplest two-state model, with the values of m_{k_f} and m_{k_u} fixed at 1.66 M^{-1} and 0.75 M^{-1}, respectively. To account for the curvature in the unfolding arm of the chevron plots, the other pathway (pathway 2) was assumed to follow the sequential transition states model, as in the described above wild-type pathway, with the values of k_{I^*-D}, m_{D-I^*}, m_{I^*-D}, m_{I^*-N} and m_{N-I^*} fixed at 1×10^4 s^{-1}, 0.80 M^{-1}, 0 M^{-1}, 1.30 ± 0.04 M^{-1} and 0.32 ± 0.03 M^{-1}, respectively. All other microscopic rate constants were allowed to vary freely. Data for E-G5[2]-G587A, E-G5[2]-G584A-E500W-E532C^{IAEDANS}, E-G5[2]-G626A and E-G5[2]-Y625W-P571A were included in the global fitting, but did not converge.

Table S7. Data collection and refinement statistics.

	E-G5^2-Y625W
Data collection *	
PDB deposition code	5DBL
Space group	*C*2
Cell dimensions	
a, *b*, *c*; Å	69.1, 35.0, 69.2
β; °	104.9
Resolution, Å	33.4—1.6 (1.63—1.60)
R_{pim}, %	4.9 (60.0)
$CC_{1/2}$§, %	99.9 (77.1)
I/σ*I*	11.1 (1.6)
Completeness, %	99.0 (98.5)
Redundancy	3.2 (3.2)
Refinement	
Resolution, Å	33.4—1.6
No. of reflections	
Working set	20011
Test set	1,074
R_{work}/R_{free}	17.4/20.7
No. of atoms	
Protein	1060
Water	298
B-factors	
Protein	22
Water	34
rmsd from ideality	
Bond lengths, Å	0.006
Bond angles, °	0.992
Ramachandran angles	
Favored regions, %	100
Outliers, %	0

* Values in parentheses are for the highest resolution shell.

§ $CC_{1/2}$ is the half-data-set correlation coefficient.

We acknowledge Johan Turkenburg and Sam Hart for assistance with crystal testing and data collection. The authors would also like to thank Diamond Light Source for beamtime (proposal mx7864), and the staff of beamline I02 for assistance with crystal testing and data collection.

Supplementary References

1. Gruszka DT, *et al.* (2015) Cooperative folding of intrinsically disordered domains drives assembly of a strong elongated protein. *Nat Commun* 6:7271.
2. Gruszka DT, *et al.* (2012) Staphylococcal biofilm-forming protein has a contiguous rod-like structure. *Proc. Natl Acad. Sci. USA* 109(17):E1011-E1018.
3. Bachmann A & Kiefhaber T (2001) Apparent two-state tendamistat folding is a sequential process along a defined route. *J. Mol. Biol.* 306(2):375-386.
4. Karanicolas J & Brooks CL, 3rd (2003) Improved Go-like models demonstrate the robustness of protein folding mechanisms towards non-native interactions. *J Mol Biol* 334(2):309-325.
5. Gruszka DT, *et al.* (2012) Staphylococcal biofilm-forming protein has a contiguous rod-like structure. *Proc. Natl Acad. Sci. USA* 109(17):E1011-E1018.
6. Newman J, *et al.* (2005) Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy. *Acta crystallographica. Section D, Biological crystallography* 61(Pt 10):1426-1431.
7. Kabsch W (2010) XDS. *Acta Crystallogr. D* 66(Pt 2):125-132.
8. Evans PR & Murshudov GN (2013) How good are my data and what is the resolution? *Acta crystallographica. Section D, Biol Crystallogr* 69(Pt 7):1204-1214.
9. McCoy AJ, *et al.* (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* 40(Pt 4):658-674.
10. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta crystallographica. Section D, Biol Crystallogr* 66(Pt 4):486-501.
11. Adams PD, *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biol Crystallogr* 66(Pt 2):213-221.
12. Krissinel E & Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D* 60(Pt 12 Pt 1):2256-2268.
13. McNicholas S, Potterton E, Wilson KS, & Noble ME (2011) Presenting your structures: the CCP4mg molecular-graphics software. *Acta crystallographica. D, Biological Crystallogr* 67(Pt 4):386-394.
14. Plaxco KW, Simons KT, & Baker D (1998) Contact order, transition state placement and the refolding rates of single domain proteins. *J Mol Biol* 277(4):985-994.